

DOCUMENT RESUME

ED 315 447

TM 014 473

TITLE Maryland Functional Reading Test-II: Fall 1987
Administration. Report R-87. Technical Report.

INSTITUTION Maryland State Dept. of Education, Baltimore. Div. of
Instruction.

PUB DATE Aug 88

NOTE 18p.

PUB TYPE Statistical Data (110) -- Reports -
Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Achievement Tests; *Grade 9; *Graduation
Requirements; High Schools; *High School Students;
Public Schools; Reading Achievement; Reading Skills;
*Reading Tests; *State Programs; Testing Programs;
*Test Results

IDENTIFIERS *Maryland Functional Reading Test

ABSTRACT

The Maryland Functional Reading Test, Level II (MFRT-II) assesses Maryland students' mastery of functional reading skills that are described in the Declared Competencies Index published by the Maryland State Department of Education in 1980. The test is used to certify students for high school graduation and to identify those in need of assistance to improve their reading skills. Forms 160 through 169 of the MFRT-II were administered to 44,032 students in Maryland public schools in grades 9 through 12 during the period October 19 through November 6, 1987. Most were ninth graders taking the test for the first time; students in grades 10 through 12 who took the test were transfers or had failed the test previously. Separate sections of this paper describe and report on the development procedures, technical characteristics, and student performance results of the 1987 edition of the MFRT-II. More detailed information is available in the technical manual. The percentage of students who passed on the first attempt stayed essentially the same as in 1986, at approximately 93% (versus 92.2%). The mean scale score in 1987 of 376.10 was about the same as that in 1986 (376.46). Nine tables present test results and passing percentages. (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

TECHNICAL REPORT

MARYLAND FUNCTIONAL READING TEST-II FALL 1987 ADMINISTRATION

**REPORT R-87
AUGUST 1988**

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

G. A. CRENSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."



**Maryland State Department of Education
Division of Instruction
Program Assessment, Evaluation,
and Instructional Support Branch
200 W. Baltimore Street
Baltimore, Maryland 21201
(301) 333-2369**

MFTP

**MARYLAND FUNCTIONAL
TESTING PROGRAM**

A COMPONENT OF PROJECT BASIC FOR:

- * ASSESSMENT OF BASIC AND CITIZENSHIP SKILLS
- * IDENTIFICATION FOR APPROPRIATE ASSISTANCE
- * CERTIFICATION FOR GRADUATION

Technical Report

MARYLAND FUNCTIONAL READING TEST-II Fall 1987 Administration

Maryland State Department of Education
Division of Instruction
Program Assessment, Evaluation, and Instructional Support Branch

August 1988

The purpose of the Maryland Functional Reading Test-II is threefold. First, the test has been designed to assess Maryland students' mastery of functional reading skills that are described in the Declared Competencies Index (DCI) published by the Maryland State Department of Education (MSDE) in 1980. In addition, the test is used to certify students for high school graduation. Finally, the test is used to identify students who should participate in appropriate assistance classes in order to improve their reading skills.

Forms 160-169 of the 1987 edition of the Maryland Functional Reading Test Level II (MFRT-II) were administered in Maryland public schools to students in grades 9-12 during the period October 19-November 6, 1987. Students in all of the state's 24 school districts took the MFRT-II for the first time in ninth grade. Students in grades 10, 11, and 12 who took the 1987 MFRT-II include those who failed the test in previous years as well as transfer students. The graduating class of 1982 was the first class required to pass the MFRT-II in order to receive a Maryland State High School diploma.

The remainder of this report contains separate sections describing and reporting the development procedures, technical characteristics, and student performance results of the 1986 edition of the MFRT-II. For more detailed information the reader is referred to the Maryland Functional Testing Program (MFTP) Technical Manual currently in production.

DEVELOPMENT

Declared Competencies Index and the MFRT-II

All Maryland public schools are required to teach the functional skills contained in the state board endorsed Declared Competencies Index (DCI). The DCI provides the basis for the test specifications as well as for the objectives assessed by the MFRT-II. In June 1982, staff from the Division of Instruction (DOI) and Project Basic produced the Project Basic Instructional Guide, Volume V: Functional Reading (MSDE, 1982). The document was designed to help teachers and supervisors develop

instructional activities to prepare students for the MFRT-II and describes the test items found in each of the test's five domains.

Test Specifications

The Test Specifications (MSDE, 1982) for the MFRT-II define the eligible content for the five domains of the MFRT-II that include: following directions, locating information, main idea, using details, and understanding forms. The test specifications describe the type of items used in each domain and provide a guide to item writers in generating new items for field testing and use on subsequent forms of the MFRT-II. This document is periodically reviewed and revised according to the general procedures described more fully in the MFTP Technical Manual. The test specifications committee for the MFRT-II did not meet prior to the development of the 1987 test, as no revisions were proposed or made to the test specifications during that time.

Item Development

The item writing committee for the MFRT-II is composed of educators from Maryland's public schools appointed by the State Superintendent of Schools. When necessary, the item writing committee convenes to generate new items to assess the objectives in each of the five domains in the test. Items are carefully checked for readability, item/objective match, and conformity with test specifications. They are reviewed, revised, and field tested prior to inclusion in the MFRT-II item bank. Although no new items were generated by the committee for the 1987 MFRT-II, several existing items from the item bank were revised, reviewed, and field tested on the 1987 MFRT-II.

Item Review

In order to be included on an edition of a Maryland Functional test, items must pass rigorous content, bias, psychometric, and editorial reviews that are described in detail in the MFTP Technical Manual. The operational and field test items included in the 1987 MFRT-II were submitted to these reviews.

Test Forms

Each year, ten forms of the MFRT-II are administered. All ten forms have the same 60 operational items that count toward a student's total score on the test. Each form of the test also includes ten field test items. This format allows the Program Assessment, Evaluation, and Instructional Support Branch (PAEISB) to gather field test data on a total of 100 items each year. For more detailed information on field test procedures and the purposes of field testing, consult the MFTP Technical Manual.

TECHNICAL CHARACTERISTICS

Reliability

The reliability of the MFRT-II is evaluated using two statistical indicators: internal consistency and classification consistency coefficients. Criterion-referenced tests (CRTs) like the MFRT-II are used to infer the degree to which students have mastered a domain of knowledge and skill. To make valid inferences from the MFRT-II to the domain of reading knowledge and skill defined by the DCI, items must measure only this knowledge and skill. Internal consistency reliability coefficients provide a measure of the extent to which the items on a test are homogeneous, or measure the same thing, thereby providing one way of describing test reliability. One internal consistency reliability coefficient is the Kuder-Richardson formula number 20, or KR20 (see, for example, Crocker & Algina, 1986; Gronlund, 1982.)

The MFRT-II is accompanied by a passing score adopted by the Maryland State Board of Education in the early 1970s. This standard corresponds to answering approximately 62% of the operational items correctly on the test. Students whose MFRT-II score is above the passing score are considered to have mastered this set of functional skills, and students scoring below the passing score are considered not to have mastered the content and, therefore, in need of remedial instruction. The standard for current editions of the MFRT-II (i.e., forms 110 and higher) was established by equating the MFRT-II to the original standard on Form-B of the Maryland Functional Reading Test. The reliability or consistency with which the MFRT-II classifies students into these two categories is essential for drawing valid, believable conclusions about a student's mastery of functional reading. An indicator of the consistency of this classification is an adaptation of a procedure developed by Subkoviak (in Berk, 1984, pp.278-280), and referred to here as the PG coefficient. The reliability and classification consistency of the 1987 MFRT-II are discussed in the following sections.

Internal consistency. Table 1 contains the internal consistency of the 1987 MFRT-II forms and items across forms. KR20 coefficients for the individual forms have remained at a stable and high rate over the last three years. On the 1987 test, they ranged from .82 to .83. Although the number of items in each domain is limited (ranging from 10 to 13), KR20 coefficients are satisfactory for all domains, ranging from .60 to 1.08 on the 1987 test. These coefficients have remained stable over the last three years. The KR20 coefficient for the total test is higher than for individual domains because domain coefficients are calculated on fewer items than the total test. KR20s are lower bound estimates of test reliability.

Table 1

Internal Consistency (KR20) and Classification Consistency (PG)
Coefficients

Individual test forms (70 items each form*)			
Form	No. Students	KR20	
160	4,378	.83	
161	4,348	.82	
162	4,368	.83	
163	4,402	.83	
164	4,375	.82	
165	4,361	.82	
166	4,287	.82	
167	4,298	.82	
168	4,295	.82	
169	4,330	.83	
All forms, all students (N=44,032)			
Domain	No. Items	KR20	PG
1. Following Directions	13	.70	.89
2. Locating Information	13	.72	.92
3. Main Ideas	10	1.08	.92
4. Using Details	13	.60	.93
5. Understanding Forms	11	.82	.92
Total Test		.91	.97

Note. The normal range of reliability coefficients is from 0 to 1. The KR20 of 1.08 for domain 3 is a computational artifact which can occur because of the limited variance of scores in domain 3. The KR20 for domain 3 indicates that internal consistency is very high for this domain.

Classification consistency. Table 1 indicates that the classification consistency of the MFRT-II is excellent: .97 for students who took the test in 1987. This coefficient reveals that approximately 97% of the students who took the test would be classified in the same way (as a passer or failer) on a second concurrent administration of the test. The PG of .97 indicates that only about 3 percent of the students who took the test might have been misclassified as passers or failers. It is likely that students who may have been classified incorrectly are those whose total scale score was just above or just below the passing score of 340.

Validity

The primary aim in developing, administering, scoring, and reporting the MFRT-II is to provide accurate decisions about students' possession of reading skills. The first condition necessary to assure the validity of any test is the demonstration of acceptable levels of reliability and classification consistency as described in the previous section. Evidence of the validity of the MFRT-II includes information about content validity, construct validity, and freedom from bias. The MFTP Technical Manual will include complete descriptions of the procedures used to help assure that the MFRT-II is administered according to prescribed procedures, methods used to determine content and construct validity, and procedures used to assist in determining that items are free of bias.

Equivalence of Forms

The issue of equity, or fairness, is of major importance to any test that has a designated passing score and is used to make decisions concerning high school graduation. Students who take the MFRT-II must be assured that the test they are taking is no more easy or difficult to pass than other editions or forms of the test. Thus, equivalence of test forms refers to equivalence of the difficulty of passing a particular form of the test.

Items are selected for inclusion on a given edition of the MFRT-II to help ensure the equivalence of current forms to previous editions. Statistical pre-equating analysis is performed before the test is administered to ensure that the forms are equivalent in difficulty to one another and to previous years' editions. This analysis also provides a way to adjust the passing score in order to assure comparability of the standard. The details of PAEISB's Rasch equating procedure can be found in the MFTP Technical Manual.

Table 2 shows raw scores (number of items correct) and their corresponding equated scale, domain, and percent correct scores and standard errors. Students who correctly answered 38 items on the 1987 MFRT-II achieved a scale score of 341, which is the lowest possible passing scale score in 1987. This passing raw

Table 2

Lookup Table for the Total Test

DOMAIN PROCEDURE							
NUMBER OF ITEMS = 60			MEAN DIFFICULTY = 1.115		STD. DEVIATION = .973		
RAW SCORE	DOMAIN SCORE	PERCENT OF TEST	LOGIT ABILITY	LOGIT STD. ERROR	SCALED ABILITY	SCALED STD. ERROR	DOMAIN SCORE STD. ERROR
0	0	0	-4.398	1.482	217	30	.83
1	1	1	-3.422	.988	237	20	1.40
2	2	3	-2.892	.731	251	15	1.99
3	4	5	-2.248	.597	260	12	2.36
4	5	6	-1.923	.532	267	11	2.72
5	7	8	-1.662	.486	272	10	3.01
6	8	10	-1.442	.452	276	9	3.25
7	10	11	-1.250	.425	280	8	3.47
8	11	13	-1.078	.398	283	8	3.60
9	13	15	-.923	.382	287	8	3.79
10	14	16	-.780	.369	289	7	3.95
11	16	18	-.647	.357	292	7	4.10
12	17	20	-.523	.347	295	7	4.23
13	19	21	-.405	.339	297	7	4.35
14	20	23	-.292	.331	299	7	4.45
15	21	25	-.185	.324	301	6	4.55
16	23	26	-.082	.318	303	6	4.64
17	24	28	.017	.313	305	6	4.72
18	26	30	.114	.308	307	6	4.80
19	27	31	.207	.302	309	6	4.84
20	29	33	.298	.299	311	6	4.91
21	31	35	.387	.296	313	6	4.97
22	31	36	.474	.293	314	6	5.03
23	33	38	.560	.291	316	6	5.08
24	35	40	.644	.289	318	6	5.13
25	37	41	.727	.287	320	6	5.17
26	38	43	.809	.286	321	6	5.20
27	39	45	.890	.285	323	6	5.24
28	40	46	.971	.284	324	6	5.27
29	42	48	1.051	.283	326	6	5.29
30	44	50	1.131	.283	328	6	5.32
31	45	51	1.211	.283	329	6	5.34
32	47	53	1.291	.283	331	6	5.35
33	48	54	1.371	.283	332	6	5.37
34	50	56	1.451	.284	334	6	5.38
35	52	58	1.532	.285	336	6	5.38
36	53	60	1.614	.287	337	6	5.39
37	55	61	1.697	.288	339	6	5.39
38	56	63	1.781	.290	341	6	5.39
39	57	65	1.866	.293	342	6	5.38
40	59	66	1.952	.295	344	6	5.37
41	61	68	2.041	.298	346	6	5.35
42	63	70	2.132	.302	348	6	5.32
43	64	71	2.225	.308	350	6	5.33
44	65	73	2.322	.313	351	6	5.31
45	67	75	2.422	.319	353	6	5.28
46	70	78	2.526	.326	356	7	5.24
47	71	78	2.634	.333	358	7	5.20
48	73	80	2.749	.342	360	7	5.14
49	74	81	2.870	.352	362	7	5.08
50	76	83	2.999	.364	365	7	5.00
51	78	85	3.137	.377	368	8	4.90
52	80	86	3.288	.393	371	8	4.78
53	82	88	3.458	.420	374	8	4.72
54	84	90	3.643	.447	378	9	4.58
55	86	91	3.859	.482	382	10	4.40
56	88	93	4.115	.529	387	11	4.16
57	90	95	4.436	.594	394	12	3.84
58	93	96	4.876	.729	403	15	3.52
59	96	98	5.601	.988	417	20	2.87
60	98	100	6.577	1.482	437	30	2.11

score is one point higher than the lowest possible passing raw score of 37 in 1986.

Table 3 contains raw scores, equated scale scores, domain scores, and percent correct scores and standard errors for each domain. To illustrate the use of this table, the panel for domain 1 indicates that a student who correctly answered 10 of 13 items, equivalent to a scale score of 345, can be considered to have mastered the domain. However, pass-fail status is not determined by performance within an individual domain: a student may perform poorly on a given domain yet pass the test by performing well in other domains. One reason for determining pass-fail status only at the total test level is that longer tests are generally more reliable than shorter tests. This relationship between test length and test reliability is reflected in the scale score standard errors (SEss) for domain one. A scale score of 345, which is based on 13 items in the domain, has a standard error of 14 -- higher than the standard error of 6 for a scale score of 344 based on 60 items. (see Tables 2 and 3)

Item Bias Analysis

The procedures used to detect item bias during the development phase of the MFRT-II will be described in detail in the MFRT Technical Manual. In addition to these judgmental reviews, items are subjected to statistical analysis to determine any sex and race bias after the test has been administered. An adaptation of the delta-plot method (Angoff, 1982, pp. 96-106) is used to identify items that may be biased so that they can be resubmitted for review prior to inclusion in a student's score on the MFRT-II.

Tables 4 and 5 contain the 1987 MFRT-II delta differences for males versus females and blacks versus whites, respectively. A delta difference of positive or negative 3.5 or higher indicates an item that might be biased and should be reviewed before final computation of student scores. Non-significant differences between deltas for each group are likely due to random fluctuations in sub-group performances and measurement error. Delta differences ($G1 - G2$) can be found in the far right column in Tables 4 and 5. All delta differences in both tables fall within the ± 3.5 guidelines, indicating that no item on the 1987 MFRT-II shows statistical evidence of sex or race bias.

RESULTS

Pass-rates. The percentage of students who took the MFRT-II for the first time and passed stayed essentially the same as for the 1986 test (see Table 6). The pass-rates for the past six years for first time takers as well as retested students are also presented.

Table 3

Equated Raw, Domain and Scale Scores and Standard Errors for Each Domain

DOMAIN 1 Following Directions							
NUMBER OF ITEMS = 13 MEAN DIFFICULTY = .544 STD. DEVIATION = 1.000							
RAW SCORE	DOMAIN SCORE	PERCENT OF TEST	LOGIT ABILITY	LOGIT STD. ERROR	SCALED ABILITY	SCALED STD. ERROR	DOMAIN SCORE STD. ERROR
0	1	0	-3.453	1.604	236	32	2.74
1	5	7	-2.310	1.069	259	21	4.99
2	10	15	-1.462	.806	276	16	7.09
3	16	23	-.893	.707	287	14	8.73
4	22	30	-.432	.651	296	13	9.96
5	30	38	-.024	.627	305	13	11.02
6	36	46	.361	.615	312	12	11.77
7	44	53	.737	.615	320	12	12.23
8	51	61	1.121	.626	327	13	12.39
9	60	69	1.528	.650	336	13	12.19
10	68	76	1.986	.704	345	14	11.75
11	76	84	2.548	.802	356	16	10.83
12	85	92	3.387	1.064	373	21	9.44
13	92	100	4.518	1.596	395	32	7.16

DOMAIN 2 Locating Information							
NUMBER OF ITEMS = 13 MEAN DIFFICULTY = .624 STD. DEVIATION = .814							
RAW SCORE	DOMAIN SCORE	PERCENT OF TEST	LOGIT ABILITY	LOGIT STD. ERROR	SCALED ABILITY	SCALED STD. ERROR	DOMAIN SCORE STD. ERROR
0	3	0	-3.254	1.594	240	32	5.21
1	9	7	-2.124	1.063	263	21	8.57
2	18	15	-1.288	.797	279	16	10.57
3	26	23	-.734	.696	290	14	11.51
4	34	30	-.290	.637	299	13	11.73
5	41	38	.100	.610	307	12	11.79
6	48	46	.462	.596	314	12	11.61
7	55	53	.815	.594	321	12	11.30
8	62	61	1.174	.605	328	12	10.90
9	68	69	1.555	.630	336	13	10.38
10	75	76	1.987	.686	345	14	9.84
11	82	84	2.524	.786	355	16	8.99
12	90	92	3.340	1.053	372	21	7.63
13	95	100	4.448	1.579	394	32	5.13

DOMAIN 3 Main Idea							
NUMBER OF ITEMS = 10 MEAN DIFFICULTY = 2.416 STD. DEVIATION = .535							
RAW SCORE	DOMAIN SCORE	PERCENT OF TEST	LOGIT ABILITY	LOGIT STD. ERROR	SCALED ABILITY	SCALED STD. ERROR	DOMAIN SCORE STD. ERROR
0	3	0	-1.012	1.596	285	32	5.50
1	9	10	.121	1.064	307	21	9.31
2	20	20	.956	.804	324	16	11.94
3	29	30	1.521	.702	335	14	13.31
4	38	40	1.987	.664	345	13	14.20
5	47	50	2.418	.652	353	13	14.51
6	57	60	2.848	.664	362	13	14.37
7	67	70	3.315	.702	371	14	13.66
8	77	80	3.880	.805	383	16	12.55
9	88	90	4.717	1.065	399	21	10.26
10	95	100	5.851	1.597	422	32	6.55

Table 3 (cont.)

Equated Raw, Domain and Scale Scores and Standard Errors for Each Domain

DOMAIN 4 Using Details							
NUMBER OF ITEMS = 13 MEAN DIFFICULTY = .964 STD. DEVIATION = .621							
RAW SCORE	DOMAIN SCORE	PERCENT OF TEST	LOGIT ABILITY	LOGIT STD. ERROR	SCALED ABILITY	SCALED STD. ERROR	DOMAIN SCORE STD. ERROR
0	1	0	-2.803	1.585	249	32	2.99
1	5	7	-1.688	1.057	271	21	5.38
2	11	15	-.863	.786	288	18	7.40
3	17	23	-.326	.683	298	14	8.89
4	23	30	.100	.621	307	12	9.86
5	29	38	.470	.593	314	12	10.72
6	36	46	.813	.579	321	12	11.35
7	42	53	1.145	.577	328	12	11.82
8	50	61	1.484	.589	335	12	12.16
9	57	69	1.846	.614	342	12	12.32
10	65	76	2.259	.673	350	13	12.44
11	74	84	2.779	.774	361	15	12.07
12	84	92	3.577	1.044	377	21	11.02
13	92	100	4.668	1.566	398	31	8.21

DOMAIN 5 Understanding Forms							
NUMBER OF ITEMS = 11 MEAN DIFFICULTY = 1.364 STD. DEVIATION = .576							
RAW SCORE	DOMAIN SCORE	PERCENT OF TEST	LOGIT ABILITY	LOGIT STD. ERROR	SCALED ABILITY	SCALED STD. ERROR	DOMAIN SCORE STD. ERROR
0	3	0	-2.228	1.607	260	32	5.74
1	10	9	-1.081	1.071	283	21	9.25
2	20	18	-.233	.806	300	18	11.21
3	29	27	.331	.703	312	14	12.03
4	36	36	.785	.649	321	13	12.23
5	44	45	1.190	.626	329	13	12.19
6	53	54	1.578	.623	337	12	11.96
7	59	63	1.975	.641	344	13	11.62
8	66	72	2.414	.689	353	14	11.22
9	75	81	2.853	.788	364	16	10.63
10	84	90	3.765	1.053	380	21	9.74
11	92	100	4.873	1.579	402	32	7.82

Table 4

Delta Values for Bias Analysis: Males vs. Females

OPERATIONAL SEQUENCE NUMBER	P-VALUE		Z-VALUE		DELTA		DELTA (G1-G2)
	G1	G2	G1	G2	G1	G2	
I001	.96	.97	1.75	1.88	20.00	20.52	-.52
I002	.96	.96	1.75	1.75	20.00	20.00	.00
I003	.88	.91	1.17	1.34	17.68	18.36	-.68
I004	.82	.81	.92	.88	16.68	16.52	.16
I005	.77	.82	.74	.92	15.96	16.68	-.72
I006	.74	.79	.64	.81	15.56	16.24	-.68
I007	.81	.86	.28	1.08	16.52	17.32	-.80
I008	.97	.97	1.88	1.88	20.52	20.52	.00
I009	.82	.87	.92	1.13	16.68	17.52	-.84
I010	.93	.95	1.48	1.65	18.92	19.60	-.68
I011	.89	.93	1.23	1.48	17.92	18.92	-1.00
I012	.83	.85	.95	1.04	16.80	17.16	-.36
I013	.98	.98	2.05	2.05	21.20	21.20	.00
I014	.83	.86	.95	1.08	16.80	17.32	-.52
I015	.56	.55	.15	.13	13.60	13.52	.08
I016	.61	.70	.28	.52	14.12	15.08	-.96
I017	.78	.86	.77	1.08	16.08	17.32	-1.24
I018	.70	.76	.52	.71	15.08	15.84	-.76
I019	.84	.89	.99	1.23	16.96	17.92	-.96
I020	.79	.84	.81	.99	16.24	16.96	-.72
I021	.68	.77	.47	.74	14.88	15.96	-1.08
I022	.61	.69	.28	.50	14.12	15.00	-.88
I023	.47	.51	-.08	.03	12.68	13.12	-.44
I024	.79	.88	.81	1.17	16.24	17.68	-1.44
I025	.91	.94	1.34	1.55	18.36	19.20	-.84
I026	.90	.94	1.28	1.55	18.12	19.20	-1.08
I027	.87	.92	1.13	1.41	17.52	18.64	-1.12
I028	.89	.93	1.23	1.48	17.92	18.92	-1.00
I029	.92	.94	1.41	1.55	18.64	19.20	-.56
I030	.96	.98	1.75	2.05	20.00	21.20	-1.20
I031	.96	.97	1.75	1.88	20.00	20.52	-.52
I032	.87	.89	1.13	1.23	17.52	17.92	-.40
I033	.91	.94	1.34	1.55	18.36	19.20	-.84
I034	.95	.97	1.65	1.88	19.60	20.52	-.92
I035	.86	.89	1.08	1.23	17.32	17.92	-.60
I036	.78	.81	.77	.88	16.08	16.52	-.44
I037	.90	.95	1.28	1.65	18.12	19.60	-1.48
I038	.89	.94	1.23	1.55	17.92	19.20	-1.28
I039	.95	.98	1.65	2.05	19.60	21.20	-1.60
I040	.80	.87	.84	1.13	16.36	17.52	-1.16
I041	.87	.90	1.13	1.28	17.52	18.12	-.60
I042	.78	.82	.77	.92	16.08	16.68	-.60
I043	.85	.91	1.04	1.34	17.16	18.36	-1.20
I044	.85	.89	1.04	1.23	17.16	17.92	-.76
I045	.85	.90	1.04	1.28	17.16	18.12	-.96
I046	.82	.84	.92	.99	16.68	16.96	-.28
I047	.90	.94	1.28	1.55	18.12	19.20	-1.08
I048	.92	.95	1.41	1.65	18.64	19.60	-.96
I049	.86	.92	1.08	1.41	17.32	18.64	-1.32
I050	.90	.95	1.28	1.65	18.12	19.60	-1.48
I051	.95	.98	1.65	2.05	19.60	21.20	-1.60
I052	.75	.81	.67	.88	15.68	16.52	-.84
I053	.82	.88	.92	1.17	16.68	17.68	-1.00
I054	.83	.90	.95	1.28	16.80	18.12	-1.32
I055	.77	.80	.74	.84	15.96	16.36	-.40
I056	.75	.81	.67	.88	15.68	16.52	-.84
I057	.83	.89	.95	1.23	16.80	17.92	-1.12
I058	.77	.86	.74	1.08	15.96	17.32	-1.36
I059	.76	.83	.71	.95	15.84	16.80	-.96
I060	.82	.89	.92	1.23	16.68	17.92	-1.24

TOTAL RECORDS--GROUP 1 = 22507: Males
 TOTAL RECORDS--GROUP 2 = 21510: Females

NOTE. Operational sequence numbers do not correspond to item numbers in actual test booklets due to the removal of field test items.

Table 5

Delta Values for Bias Analysis: Whites vs. Blacks

OPERATIONAL		P-VALUE		Z-VALUE		DELTA		DELTA
SEQUENCE	NUMBER	G1	G2	G1	G2	G1	G2	(G1-G2)
I001		.98	.95	2.05	1.85	21.20	19.60	1.60
I002		.98	.93	2.05	1.48	21.20	18.92	2.28
I003		.91	.88	1.34	1.17	18.36	17.68	.68
I004		.87	.72	1.13	.58	17.52	15.32	2.20
I005		.86	.70	1.08	.52	17.32	15.08	2.24
I006		.84	.65	.99	.39	16.96	14.56	2.40
I007		.88	.76	1.17	.71	17.68	15.84	1.84
I008		.98	.96	2.05	1.75	21.20	20.00	1.20
I009		.87	.80	1.13	.84	17.52	16.36	1.16
I010		.96	.91	1.75	1.34	20.00	18.36	1.64
I011		.94	.86	1.55	1.08	19.20	17.32	1.88
I012		.88	.78	1.17	.77	17.68	16.08	1.60
I013		.99	.96	2.33	1.75	22.32	20.00	2.32
I014		.89	.79	1.23	.81	17.92	16.24	1.68
I015		.61	.48	.28	-.05	14.12	12.80	1.32
I016		.70	.59	.52	.23	15.08	13.92	1.16
I017		.85	.77	1.04	.74	17.16	15.96	1.20
I018		.75	.72	.67	.58	15.68	15.32	.36
I019		.90	.81	1.28	.88	18.12	16.52	1.60
I020		.88	.75	1.08	.67	17.32	15.68	1.64
I021		.77	.65	.74	.39	15.96	14.56	1.40
I022		.70	.57	.52	.18	15.08	13.72	1.36
I023		.52	.45	.05	-.13	13.20	12.48	.72
I024		.89	.76	1.23	.71	17.92	15.84	2.08
I025		.95	.89	1.65	1.23	19.60	17.92	1.68
I026		.94	.89	1.55	1.23	19.20	17.92	1.28
I027		.92	.84	1.41	.99	18.64	16.96	1.68
I028		.93	.88	1.48	1.17	18.92	17.68	1.24
I029		.93	.92	1.48	1.41	18.92	18.64	.28
I030		.98	.95	2.05	1.65	21.20	19.60	1.60
I031		.98	.96	2.05	1.75	21.20	20.00	1.20
I032		.88	.87	1.17	1.13	17.68	17.52	.16
I033		.94	.91	1.55	1.34	19.20	18.36	.84
I034		.98	.94	2.05	1.55	21.20	19.20	2.00
I035		.89	.86	1.23	1.08	17.92	17.32	.60
I036		.82	.76	.92	.71	16.68	15.84	.84
I037		.95	.89	1.65	1.23	19.60	17.92	1.68
I038		.92	.90	1.41	1.28	18.64	18.12	.52
I039		.98	.95	2.05	1.65	21.20	19.60	1.60
I040		.89	.74	1.23	.64	17.92	15.96	2.36
I041		.91	.84	1.34	.99	18.36	16.96	1.40
I042		.85	.72	1.04	.58	17.16	15.32	1.84
I043		.91	.85	1.34	1.04	18.36	17.16	1.20
I044		.90	.82	1.28	.92	18.12	16.68	1.44
I045		.91	.82	1.34	.92	18.36	16.68	1.68
I046		.85	.81	1.04	.88	17.16	16.52	.64
I047		.95	.88	1.65	1.17	19.60	17.68	1.92
I048		.95	.91	1.65	1.34	19.60	18.36	1.24
I049		.92	.84	1.41	.99	18.64	16.96	1.68
I050		.95	.88	1.65	1.17	19.60	17.68	1.92
I051		.98	.95	2.05	1.65	21.20	19.60	1.60
I052		.82	.71	.92	.55	16.68	15.20	1.48
I053		.80	.78	1.23	.77	17.92	16.08	1.84
I054		.89	.81	1.23	.88	17.92	16.52	1.40
I055		.79	.78	.81	.77	16.24	16.08	.16
I056		.82	.72	.92	.58	16.68	15.32	1.36
I057		.89	.81	1.23	.88	17.92	16.52	1.40
I058		.84	.76	.99	.71	16.96	15.84	1.12
I059		.84	.74	.99	.64	16.96	15.56	1.40
I060		.91	.77	1.34	.74	18.36	15.96	2.40

TOTAL RECORDS--GROUP 1 = 25086: Whites
 TOTAL RECORDS--GROUP 2 = 9842: Blacks

NOTE. Operational sequence numbers do not correspond to item numbers in actual test booklets due to the removal of field test items.

Table 6

Pass-Rates for the MFRT-II: 1982-1987

Grade	Percent Passing					
	School Year					
	1982-83	1983-84	1984-85	1985-86	1986-87	1987-88
9	89.1	93.1	93.1	92.2	92.2	93.1
10	67.1	NA	78.6	76.8	74.5	78.0
11	74.7	NA	84.2	81.3	83.0	84.6
12	78.0	NA	87.7	84.0	85.1	88.1

Note. Percentages include fall administration only.

NA. Not available.

Response Proportions. Table 7 contains the response proportions and p-values for each operational item on the MFRT-II. The response proportion indicates the proportion of students selecting each possible response for each item on the MFRT-II. Response proportions are used to evaluate the quality of the incorrect responses (or distractors), and can be useful in identifying confusing or miskeyed items. The p-value indicates the percentage of students who answered an item correctly. The p-values for the 1987 MFRT-II range from .49 (operational sequence number 23, actual item number 26) to .98 (operational sequence number 13, actual item number 14). Experience and convention has defined the acceptable range of p-values as .30 to .70 for norm-referenced tests. However, this convention applies less strictly to criterion-referenced tests that may include items outside of this range in order to ensure adequate coverage of relevant content areas. In particular, it is appropriate for basic skills tests like the MFRT-II to contain a number of "easier" items -- with p-values greater than .70 -- to ensure accurate measurement of low-achieving students and meaningful diagnostic domain scores.

Summary Results. The mean raw, domain, and scale scores and their standard deviations for the 1987 MFRT-II are presented in Table 8. The mean scale score of 376.10 remained essentially the same as 1986's mean scale score of 376.46. Finally, Table 9

Table 7

Percentages of Students Who Answered Each Item Correctly (P-Values) and
Response Proportions

SEQUENCE NUMBER	A	RESPONSE B	PROPORTIONS C	D	OTHER	P-VALUE
I001	.97	.02	.01	.01	.00	.97
I002	.01	.96	.02	.00	.00	.96
I003	.01	.90	.04	.05	.00	.90
I004	.12	.81	.04	.02	.00	.81
I005	.10	.01	.80	.09	.00	.80
I006	.77	.05	.13	.05	.00	.77
I007	.09	.04	.84	.03	.00	.84
I008	.97	.01	.01	.01	.00	.97
I009	.06	.07	.02	.84	.00	.84
I010	.02	.02	.94	.02	.00	.94
I011	.02	.01	.06	.91	.00	.91
I012	.06	.84	.04	.06	.00	.84
I013	.01	.88	.01	.01	.00	.88
I014	.10	.03	.84	.03	.00	.84
I015	.21	.55	.02	.21	.00	.55
I016	.09	.23	.03	.66	.00	.66
I017	.07	.05	.82	.06	.00	.82
I018	.04	.73	.20	.02	.00	.73
I019	.03	.05	.86	.06	.00	.86
I020	.02	.81	.10	.06	.00	.81
I021	.14	.08	.05	.73	.00	.73
I022	.65	.06	.21	.08	.00	.65
I023	.09	.36	.49	.05	.00	.49
I024	.08	.02	.83	.07	.00	.83
I025	.02	.93	.01	.04	.00	.93
I026	.01	.04	.03	.92	.00	.92
I027	.05	.89	.04	.02	.00	.89
I028	.07	.91	.01	.01	.00	.91
I029	.05	.01	.01	.93	.00	.93
I030	.01	.97	.02	.00	.00	.97
I031	.97	.01	.02	.00	.00	.97
I032	.01	.05	.88	.06	.00	.88
I033	.01	.03	.03	.93	.00	.93
I034	.01	.01	.96	.01	.00	.96
I035	.10	.88	.02	.01	.00	.88
I036	.02	.15	.03	.79	.00	.79
I037	.03	.02	.92	.02	.00	.92
I038	.01	.03	.04	.91	.00	.91
I039	.01	.97	.01	.01	.00	.97
I040	.03	.84	.04	.09	.00	.84
I041	.06	.03	.02	.88	.00	.88
I042	.80	.03	.10	.07	.00	.80
I043	.88	.09	.02	.01	.00	.88
I044	.05	.07	.87	.01	.00	.87
I045	.06	.87	.03	.02	.00	.87
I046	.83	.13	.02	.02	.00	.83
I047	.03	.01	.03	.92	.00	.92
I048	.84	.04	.02	.01	.00	.84
I049	.89	.03	.02	.06	.00	.89
I050	.03	.02	.02	.92	.00	.92
I051	.01	.97	.01	.01	.00	.97
I052	.78	.02	.10	.09	.00	.78
I053	.10	.85	.03	.02	.00	.85
I054	.01	.10	.02	.86	.00	.86
I055	.02	.02	.79	.17	.01	.79
I056	.01	.10	.10	.78	.01	.78
I057	.01	.07	.88	.05	.01	.88
I058	.01	.07	.81	.10	.01	.81
I059	.09	.80	.03	.07	.01	.80
I060	.07	.06	.85	.01	.01	.85

TOTAL RECORDS = 44032

NOTE. Operational sequence numbers do not correspond to item numbers in actual test booklets due to the removal of field test items.

presents the distribution of frequencies of each score point for the 1987 MFRT-II.

Table 8

Mean Raw, Domain, and Scale Scores and Standard Deviations
(in parentheses): Grade 9 (N=44,032)

Domain	Maximum Raw Score	Raw Score	Domain Score	Scale Score
Following Directions	13	11.5 (1.9)	80.4 (14.8)	370.6 (24.8)
Locating Information	13	11.8 (1.8)	87.4 (12.1)	373.4 (22.9)
Main Ideas	10	7.2 (2.2)	68.8 (21.4)	377.6 (26.4)
Using Details	13	11.5 (2.2)	76.7 (17.1)	375.7 (25.4)
Understanding Forms	11	9.3 (2.0)	77.7 (16.5)	374.4 (25.8)
All Domains	60	51.24 (8.19)	-- --	376.10 (24.45)

Table 9

Frequency Distribution of Total Scores

RAW SCORE	SCALE SCORE	PERCENT SCORE	FREQUENCY	FREQUENCY PERCENTAGE	CUMULATIVE FREQUENCY	CUMULATIVE PERCENTAGE
0	217	.0	1	.0	1	.0
1	237	1.7	0	.0	1	.0
2	251	3.3	0	.0	1	.0
3	260	5.0	1	.0	2	.0
4	267	6.7	2	.0	4	.0
5	272	8.3	1	.0	5	.0
6	276	10.0	3	.0	8	.0
7	280	11.7	7	.0	15	.0
8	283	13.3	9	.0	24	.1
9	287	15.0	8	.0	32	.1
10	289	16.7	10	.0	42	.1
11	292	18.3	23	.1	65	.1
12	295	20.0	23	.1	88	.2
13	297	21.7	27	.1	115	.3
14	299	23.3	49	.1	164	.4
15	301	25.0	48	.1	212	.5
16	303	26.7	37	.1	249	.6
17	305	28.3	80	.2	329	.7
18	307	30.0	65	.1	394	.9
19	309	31.7	68	.2	462	1.0
20	311	33.3	88	.2	550	1.2
21	313	35.0	80	.2	630	1.4
22	314	36.7	86	.2	716	1.6
23	316	38.3	91	.2	807	1.8
24	318	40.0	99	.2	906	2.1
25	320	41.7	94	.2	1000	2.3
26	321	43.3	126	.3	1126	2.6
27	323	45.0	120	.3	1246	2.8
28	324	46.7	135	.3	1381	3.1
29	326	48.3	138	.3	1519	3.4
30	328	50.0	140	.3	1659	3.8
31	329	51.7	159	.4	1818	4.1
32	331	53.3	157	.4	1975	4.5
33	332	55.0	184	.4	2159	4.9
34	334	56.7	177	.4	2336	5.3
35	336	58.3	203	.5	2539	5.8
36	337	60.0	225	.5	2764	6.3
37	339	61.7	277	.6	3041	6.9
38	341	63.3	292	.7	3333	7.6
39	342	65.0	344	.8	3677	8.4
40	344	66.7	377	.9	4054	9.2
41	346	68.3	447	1.0	4501	10.2
42	348	70.0	542	1.2	5043	11.5
43	350	71.7	589	1.3	5632	12.8
44	351	73.3	727	1.7	6359	14.4
45	353	75.0	845	1.9	7204	16.4
46	356	76.7	992	2.3	8196	18.8
47	358	78.3	1146	2.6	9342	21.2
48	360	80.0	1288	2.9	10630	24.1
49	362	81.7	1477	3.4	12107	27.5
50	365	83.3	1773	4.0	13880	31.5
51	368	85.0	2088	4.7	15968	36.3
52	371	86.7	2457	5.6	18425	41.8
53	374	88.3	2758	6.3	21183	48.1
54	378	90.0	3371	7.7	24554	55.8
55	382	91.7	3792	8.6	28346	64.4
56	387	93.3	4201	9.5	32547	73.9
57	394	95.0	4216	9.6	36763	83.5
58	403	96.7	3701	8.4	40464	91.9
59	417	98.3	2631	6.0	43095	97.9
60	437	100.0	937	2.1	44032	100.0

NUMBER OF RECORDS = 44032

References

- Angoff, W.H. (1982). Use of difficulty and discrimination indices for detecting item bias. In R. A. Berk (Ed.), Handbook of methods for detecting item bias. Baltimore: The Johns Hopkins University Press.
- Chafin, A. E. (1983). Setting a standard for standard setting. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Canada.
- Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. NY: CBS College Publishing.
- Gronlund, N. E. (1982). Constructing achievement tests (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Maryland State Department of Education (1980). Declared competencies index. Baltimore, MD: Author.
- Maryland State Department of Education. (1981). Project basic instructional guide. Volume VI: Functional reading. Baltimore, MD: Author.
- Subkoviak, M. (1984) Estimating the Reliability of Mastery-Nonmastery Classifications. In R. A. Berk (Ed.), A guide to criterion-referenced test construction. Baltimore: The Johns Hopkins University Press.

LWB16001